Solutions to Review Problems for Exam #3

1. Let X have a Gamma distribution with parameters $\alpha = 4$ and $\beta = \theta > 0$.

   (a) Find the Fisher information $I(\theta)$.

   Solution: The pdf of X is given by
   $$f(x \mid \theta) = \frac{1}{\Gamma(4)\theta^4}\; x^3\; e^{-x/\theta} \qquad \text{for } 0 < x < \infty$$
   and zero elsewhere, where $\Gamma(4) = \Gamma(3{+}1) = 3! = 6$. Consequently,
   $$\ln f(x \mid \theta) = -4\ln\theta - \frac{x}{\theta} + 3\ln x - \ln 6.$$
   It then follows that
   $$\frac{\partial}{\partial\theta}(\ln f(x \mid \theta)) = -\frac{4}{\theta} + \frac{x}{\theta^2},$$
   and
   $$\frac{\partial^2}{\partial\theta^2}(\ln f(x \mid \theta)) = \frac{4}{\theta^2} - \frac{2x}{\theta^3}.$$
   It then follows that the Fisher information is
   $$I(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}(\ln f(x \mid \theta))\right) = -\frac{4}{\theta^2} + \frac{2E(X)}{\theta^3},$$
   where $E(X) = 4\theta$. Thus,
   $$I(\theta) = \frac{4}{\theta^2}.$$
   □

   (b) Let $X_1, X_2, \ldots, X_n$ be a random sample from a Gamma$(4, \theta)$ distribution. Find the MLE for θ and show that it is an efficient estimator.

   Solution: The likelihood function in this case is
   $$L(\theta \mid x_1, x_2, \ldots, x_n) = \frac{1}{6^n\theta^{4n}}\;(x_1 \cdot x_2 \cdots x_n)^3\; e^{-y/\theta},$$
   where $y = \displaystyle\sum_{i=1}^{n} x_i$. We then have that
   $$\ell(\theta) = \ln L(\theta \mid x_1, x_2, \ldots, x_n) = -4n\ln\theta - \frac{y}{\theta} + 3\ln(x_1 \cdot x_2 \cdots x_n) - n\ln 6.$$

To find the MLE for θ, we maximize the function $\ell(\theta)$. Thus, we first computet the derivatives

$$\ell'(\theta) = -\frac{4n}{\theta} + \frac{y}{\theta^2},$$

and

$$\ell'(\theta) = \frac{4n}{\theta^2} - \frac{2y}{\theta^3}.$$

Thus, $\widehat{\theta} = \dfrac{y}{4n}$ is a critical point of ℓ with

$$\ell'(\widehat{\theta}) = \frac{4n}{\widehat{\theta}^2} - \frac{8n\widehat{\theta}}{\widehat{\theta}^3} = -\frac{4n}{\widehat{\theta}^2} < 0.$$

Thus, $\widehat{\theta}$ is the MLE for θ.
Observe that

$$E(\widehat{\theta}) = \frac{1}{4n}E(Y) = \frac{1}{4n}\sum_{i=1}^{n} E(X_i) = \frac{1}{4n}\sum_{i=1}^{n} 4\theta = \theta.$$

Thus, $\widehat{\theta}$ is an unbiased estimator of θ.
To see if $\widehat{\theta}$ is efficient, we compute the variance of $\widehat{\theta}$:

$$\begin{aligned}
\mathrm{var}(\widehat{\theta}) &= \frac{1}{4^2n^2}\mathrm{var}(Y) \\
&= \frac{1}{4^2n^2}\sum_{i=1}^{n} \mathrm{var}(X_i) \\
&= \frac{1}{4^2n^2}\sum_{i=1}^{n} 4\theta^2 \\
&= \frac{\theta^2}{4n}.
\end{aligned}$$

Observe that

$$\mathrm{var}(\widehat{\theta}) = \frac{1}{n(4/\theta^2)} = \frac{1}{nI(\theta)},$$

by the result from part (a), which is the Crámer–Rao lower bound. Hence, $\widehat{\theta}$ is an efficient estimator of θ. □

2. Let X have a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = \theta > 0$.

   (a) Find the Fisher information $I(\theta)$.

   Solution: The pdf of X in this case is

$$f(x \mid \theta) = \frac{1}{\sqrt{2\pi}\ \sqrt{\theta}}\ e^{-x^2/2\theta}, \quad \text{for} \quad x \in \mathbb{R}.$$

   We then have that

$$\ln f(x \mid \theta) = -\frac{1}{2}\ln\theta - \frac{x^2}{2\theta} - \frac{1}{2}\ln(2\pi).$$

   Thus,

$$\frac{\partial}{\partial\theta}(\ln f(x \mid \theta)) = -\frac{1}{2\theta} + \frac{x^2}{2\theta^2},$$

   and

$$\frac{\partial^2}{\partial\theta^2}(\ln f(x \mid \theta)) = \frac{1}{2\theta^2} - \frac{x^2}{\theta^3}.$$

   We then have that the Fisher information is

$$I(\theta) = -E\left(\frac{\partial^2}{\partial\theta^2}(\ln f(x \mid \theta))\right) = -\frac{1}{2\theta^2} + \frac{E(X^2)}{\theta^3},$$

   where $E(X^2) = \text{var}(X) = \theta$. Consequently,

$$I(\theta) = \frac{1}{2\theta^2}.$$

   □

   (b) Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal$(0, \theta)$ distribution. Find the MLE for θ and show that it is an efficient estimator.

   Solution: The likelihood function is

$$L(\theta \mid x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\ \theta^{n/2}}\ e^{-\sum_{i=1}^n x_i^2/2\theta}.$$

   To find an MLE for θ, we need to maximize the function

$$\ell(\theta) = \ln L(\theta \mid x_1, x_2, \ldots, x_n) = -\frac{n}{2}\ln\theta - \frac{1}{2\theta}\sum_{i=1}^n x_i^2 - \frac{n}{2}\ln(2\pi).$$

over $\theta > 0$. In order to do this, we compute the derivatives

$$\ell'(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n} x_i^2$$

and

$$\ell''(\theta) = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n} x_i^2.$$

We then have that

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

is a critical point of ℓ with

$$\ell''(\widehat{\theta}) = \frac{n}{2\widehat{\theta}^2} - \frac{1}{\widehat{\theta}^3}\sum_{i=1}^{n} x_i^2 = -\frac{n}{2\widehat{\theta}^2},$$

from which we conclude that $\widehat{\theta}$ is the MLE of θ.
The expectation of $\widehat{\theta}$ is

$$E(\widehat{\theta}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^2) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{var}(X_i) = \theta,$$

since the X_is are iid normal$(0, \theta)$ random variables. It then follows that $\widehat{\theta}$ is an unbiased estimator.
To compute the variance of $\widehat{\theta}$,

$$\mathrm{var}(\widehat{\theta}) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{var}(X_i^2),$$

observe that $\dfrac{1}{\theta}X_i^2$ has a $\chi^2(1)$ distribution. We then have that

$$\mathrm{var}\left(\frac{1}{\theta}X_i^2\right) = 2, \quad \text{for } i = 1, 2, \ldots, n.$$

Consequently,

$$\mathrm{var}\left(X_i^2\right) = \theta^2\ \mathrm{var}\left(\frac{1}{\theta}X_i^2\right) = 2\theta^2, \quad \text{for } i = 1, 2, \ldots, n,$$

and, therefore,

$$\text{var}(\widehat{\theta}) = \frac{1}{n^2} \cdot n(2\theta^2) = \frac{2\theta^2}{n}.$$

Thus, we can write

$$\text{var}(\widehat{\theta}) = \frac{1}{n(1/2\theta^2)} = \frac{1}{nI(\theta)},$$

by the result from part (a). This is the Crámer–Rao lower bound. Hence, $\widehat{\theta}$ is an efficient estimator of θ. $\square$

3. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a uniform distribution over the interval $[0, \theta]$ for some parameter $\theta > 0$.

   Show that $W = 2\overline{X}_n$ is an unbiased estimator of θ and determine its efficiency.

***Solution**:* Evaluate the expectation of W:

$$E(W) = E(2\overline{X}_n) = 2E(\overline{X}_n) = 2 \cdot \frac{\theta}{2} = \theta.$$

Thus, W is unbiased. Next, compute the variance of W:

$$\text{var}(W) = \text{var}(2\overline{X}_n) = 4\text{var}(\overline{X}_n) = 4 \cdot \frac{\theta^2/12}{n} = \frac{\theta^2}{3n}.$$

We have seen in class that the Crámer–Rao inequality does not apply to samples from the uniform$[0, \theta]$ distribution. However, we can compare the variance of W with that of another unbiased estimator, namely

$$W_2 = \frac{n+1}{n} \max\{X_1, X_2, \ldots, X_n\},$$

whose variance is

$$\text{var}(W_2) = \frac{1}{n(n+2)}\,\theta^2.$$

Observe that $n(n+2) > 3n$ for all $n > 1$. Consequently,

$$\text{var}(W_2) < \text{var}(W),$$

and, therefore, W is less efficient than W_2. $\square$

4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal$(0, \theta)$ distribution. We want to use the statistic
$$Y = \sum_{i=1}^{n} |X_i|$$
to estimate the standard deviation $\sqrt{\theta}$.

   (a) Let $W = cY$ for some constant c. Determine a value of c so that W is an unbiased estimator of $\sqrt{\theta}$.

   Solution: We first compute the expectation of Y:
$$E(Y) = \sum_{i=1}^{n} E\left(|X_i|\right),$$
where, for each $i = 1, 2, \ldots, n$,
$$E\left(|X_i|\right) \;=\; \int_{-\infty}^{\infty} |x| f(x \mid \theta) \ \mathrm{d}x,$$
where
$$f(x \mid \theta) = \frac{1}{\sqrt{2\pi}\ \sqrt{\theta}}\ e^{-x^2/2\theta}, \quad \text{for} \ \ x \in \mathbb{R}.$$
We then have, by the symmetry of the pdf, that
$$\begin{aligned} E\left(|X_i|\right) &= 2\int_0^{\infty} x\ \frac{1}{\sqrt{2\pi}\ \sqrt{\theta}}\ e^{-x^2/2\theta} \ \ \mathrm{d}x \\ &= \frac{2}{\sqrt{2\pi}\ \sqrt{\theta}} \int_0^{\infty} x\ e^{-x^2/2\theta} \ \ \mathrm{d}x. \end{aligned}$$
Next, make the change of variables $u = \dfrac{x^2}{2\theta}$, so that $\mathrm{d}u = \dfrac{x}{\theta}\ \mathrm{d}x$, and
$$E\left(|X_i|\right) \;=\; \frac{2\sqrt{\theta}}{\sqrt{2\pi}} \int_0^{\infty} e^{-u} \ \ \mathrm{d}u = \sqrt{\frac{2}{\pi}}\ \sqrt{\theta}.$$
We then have that
$$E(Y) = \sum_{i=1}^{n} E\left(|X_i|\right) = n \cdot \sqrt{\frac{2}{\pi}}\ \sqrt{\theta}.$$
Thus, setting $c = \dfrac{1}{n} \cdot \sqrt{\dfrac{\pi}{2}}$, we see that $W = cY$ is an unbiased estimator for $\sqrt{\theta}$. □

(b) Compute the efficiency of the estimator W found in part (a).

Solution: We first compute the variance of W:

$$\text{var}(W) = c^2 \ \text{var}(Y),$$

where

$$\text{var}(Y) = \sum_{i=1}^{n} \text{var}\left(|X_i|\right) = n \cdot \text{var}(|X|),$$

for $X \sim \text{normal}(0, \theta)$.

Thus, we need to compute $\text{var}(|X|)$:

$$\begin{aligned} \text{var}(|X|) &= E(X^2) - [E(|X|)]^2 \\ &= \text{var}(X) - \frac{2}{\pi}\,\theta \\ &= \theta - \frac{2}{\pi}\,\theta \\ &= \left(1 - \frac{2}{\pi}\right)\,\theta. \end{aligned}$$

We then have that

$$\text{var}(Y) = n \cdot \left(1 - \frac{2}{\pi}\right)\,\theta,$$

and, consequently,

$$\text{var}(W) = c^2 \ \text{var}(Y) = \frac{1}{n^2} \cdot \frac{\pi}{2} \cdot n \cdot \left(1 - \frac{2}{\pi}\right)\,\theta,$$

or

$$\text{var}(W) = \frac{1}{n}\left(\frac{\pi}{2} - 1\right)\,\theta.$$

Next, we compute the efficiency of W,

$$\text{eff}_{\theta}(W) = \frac{1/nI(\sqrt{\theta})}{\text{var}(W)}.$$

In order to compute $I(\sqrt{\theta})$, we need to write the pdf of $X \sim \text{normal}(0, \theta)$ as a function of $\sqrt{\theta}$. Setting $\sigma = \sqrt{\theta}$, we can write

$$f(x \mid \sigma) = \frac{1}{\sqrt{2\pi}\ \sigma}\ e^{-x^2/2\sigma^2}, \quad \text{for} \quad x \in \mathbb{R}.$$

We then obtain that

$$\ln f(x \mid \sigma) = -\ln\sigma - \frac{x^2}{2\sigma^2} - \frac{1}{2}\ln(2\pi),$$

so that

$$\frac{\partial}{\partial\sigma}\ln f(x \mid \eta) = -\frac{1}{\sigma} + \frac{x^2}{\sigma^3}$$

and

$$\frac{\partial^2}{\partial\sigma^2}\ln f(x \mid \eta) = \frac{1}{\sigma^2} - \frac{3x^2}{\sigma^4}.$$

Thus,

$$I(\sigma) = -E\left(\frac{\partial^2}{\partial\sigma^2}\ln f(x \mid \sigma)\right) = -\frac{1}{\sigma^2} + \frac{3E(X^2)}{\sigma^4},$$

where $E(X^2) = \text{var}(X) = \sigma^2$. Thus,

$$I(\sigma) = \frac{2}{\sigma^2} = \frac{2}{\theta}.$$

Hence,

$$\begin{aligned}
\text{eff}_\theta(W) &= \frac{1/nI(\sqrt{\theta})}{\text{var}(W)} \\
&= \frac{1/n(2/\theta)}{\text{var}(W)} \\
&= \frac{\dfrac{\theta}{2n}}{\dfrac{1}{n}\left(\dfrac{\pi}{2}-1\right)\theta} \\
&= \frac{1}{\pi-2}.
\end{aligned}$$

□

5. Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal(μ_o, θ) distribution, where μ_o is known and $\theta > 0$. Show that the LRT for

$$\text{H}_o\colon \quad \theta = \theta_o \quad \text{versus} \quad \text{H}_1\colon \quad \theta \neq \theta_o$$

may be based upon the statistic

$$W = \frac{1}{\theta_o}\sum_{i=1}^{n}(X_i - \mu_o)^2.$$

Determine the null distribution of W and give, explicitly, the rejection rule for a level α test.

Solution: The likelihood function in this case is

$$L(\theta \mid x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\ \theta^{n/2}}\ e^{-\sum_{i=1}^{n}(x_i-\mu_o)^2/2\theta}.$$

We first find the MLE for θ. In order to do this, we maximize the function

$$\ell(\theta) = \ln L(\theta \mid x_1, x_2, \ldots, x_n) = -\frac{n}{2}\ln\theta - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i - \mu_o)^2 - \frac{n}{2}\ln(2\pi),$$

for $\theta > 0$. Taking the derivatives of ℓ we obtain

$$\ell'(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \mu_o)^2,$$

and

$$\ell''(\theta) = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n}(x_i - \mu_o)^2.$$

Thus,

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_o)^2$$

is a critical point of ℓ with $\ell''(\widehat{\theta}) = -\dfrac{n}{2\widehat{\theta}^2} < 0$, and therefore it is the MLE of θ. The likelihood ratio statistic for the test of

$$\mathrm{H}_o\colon \quad \theta = \theta_o \quad \text{versus} \quad \mathrm{H}_1\colon \quad \theta \neq \theta_o$$

is

$$\Lambda(x_1, x_2, \ldots, x_n) = \frac{L(\theta_o \mid x_1, \ldots, x_n)}{L(\widehat{\theta} \mid x_1, \ldots, x_n)},$$

where

$$
\begin{aligned}
L(\theta \mid x_1, x_2, \ldots, x_n) &= \frac{1}{(2\pi)^{n/2}\ \theta_o^{n/2}}\ e^{-\sum_{i=1}^n (x_i-\mu_o)^2/2\theta_o} \\
&= \frac{1}{(2\pi)^{n/2}\ \theta_o^{n/2}}\ e^{-\frac{n}{2}\ \frac{\widehat{\theta}}{\theta_o}},
\end{aligned}
$$

and

$$
L(\widehat{\theta} \mid x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\widehat{\theta}^{n/2}}\ e^{-\frac{n}{2}}.
$$

It then follows that

$$
\Lambda(x_1, x_2, \ldots, x_n) = e^{n/2}\left(\frac{\widehat{\theta}}{\theta_o}\right)^{n/2}\ e^{-\frac{n}{2}\ \frac{\widehat{\theta}}{\theta_o}},
$$

which we can write as

$$
\Lambda(x_1, x_2, \ldots, x_n) = e^{n/2}t^{n/2}\ e^{-\frac{n}{2}\ t},
$$

where

$$
t = \frac{\widehat{\theta}}{\theta_o}.
$$

Thus, Λ is a function of t; more precisely,

$$
\Lambda(x_1, x_2, \ldots, x_n) = g(t),
$$

where $g(t) = e^{n/2}t^{n/2}e^{-tn/2}$ for $t \geqslant 0$. A sketch of the graph of g, for the case $n = 10$, is shown in Figure 1 on page 11. We then see that $g(t) \leqslant 1$ for all t, $g(0) = 0$, $\lim_{t\to\infty} g(t) = 0$, $g(t)$ is strictly increasing for $t < 1$, and strictly decreasing for $t > 0$. Thus, the LRT rejection region

$$
R\colon \quad \Lambda \leqslant c, \quad \text{for some } c \in (0,1),
$$

is equivalent to

$$
R\colon \quad t \leqslant t_1 \text{ or } t \geqslant t_2 \quad \text{for some } 0 < t_1 < 1 < t_2,
$$

or, equivalently,

$$
R\colon \quad \frac{\widehat{\theta}}{\theta_o} \leqslant t_1 \text{ or } \frac{\widehat{\theta}}{\theta_o} \geqslant t_2 \quad \text{for some } 0 < t_1 < 1 < t_2,
$$

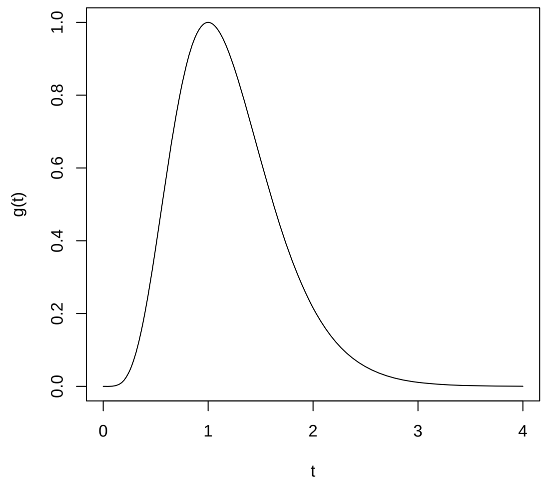

Figure 1: Sketch of graph of $g(t)$ for $0 \leqslant t \leqslant 4$

where

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_o)^2.$$

Thus the LRT may be based upon the statistic

$$W = \frac{1}{\theta_o}\sum_{i=1}^{n}(X_i - \mu_o)^2.$$

If the null hypothesis is true, then $X_1, X_2, \ldots, X_n$ are iid normal(μ_o, θ_o) random variables, and therefore

$$\frac{1}{\theta_o}(X_i - \mu_o)^2 \sim \chi^2(1) \quad \text{for } i = 1, 2, \ldots, n.$$

It then follows that, if H_o is true, then

$$W \sim \chi^2(n)$$

by the definition of the χ^2 distribution. Hence, if $a = F_W^{-1}(\alpha/2)$ and $b = F_W^{-1}(1 - (\alpha/2))$, the the LRT which rejects H_o if

$$W \leqslant a \quad \text{ or } \quad W \geqslant b,$$

has significance level α. $\square$

6. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Gamma$(4, \theta)$ distribution with $\theta > 0$.

   (a) Show that the LRT for

   $$\mathrm{H}_o\colon \quad \theta = \theta_o \quad \text{versus} \quad \mathrm{H}_1\colon \quad \theta \neq \theta_o$$

   may be based upon the statistic

   $$W = \sum_{i=1}^{n} X_i.$$

   Determine the null distribution of $2W/\theta_o$.

   Solution: The likelihood function in this case is

   $$L(\theta \mid x_1, x_2, \ldots, x_n) = \frac{1}{6^n \theta^{4n}} \, (x_1 \cdot x_2 \cdots x_n)^3 \, e^{-y/\theta},$$

   where $y = \displaystyle\sum_{i=1}^{n} x_i$. We showed in part (b) of Problem 1 in this set of review problems that $\widehat{\theta} = \dfrac{y}{4n}$ is the MLE for θ. Consequently, the likelihood ratio statistic in this case is

   $$\begin{aligned} \Lambda(x_1, x_2, \ldots, x_n) &= \frac{L(\theta_o \mid x_1, x_2, \ldots, x_n)}{L(\widehat{\theta} \mid x_1, x_2, \ldots, x_n)} \\ &= e^{4n} \left(\frac{\widehat{\theta}}{\theta_o} \right)^{4n} e^{-4n\widehat{\theta}/\theta_o}, \end{aligned}$$

   where we have used the fact that $y = 4n\widehat{\theta}$. Thus, setting

   $$t = \frac{\widehat{\theta}}{\theta_o},$$

   we see that

   $$\Lambda(x_1, x_2, \ldots, x_n) = g(t),$$

   where

   $$g(t) = t^{4n} e^{-4n(t-1)}.$$

   A sketch of the graph of g for the case $n = 5$ is shown in Figure 2 on page 13. Thus, we see from the sketch that for any $c \in (0, 1)$,

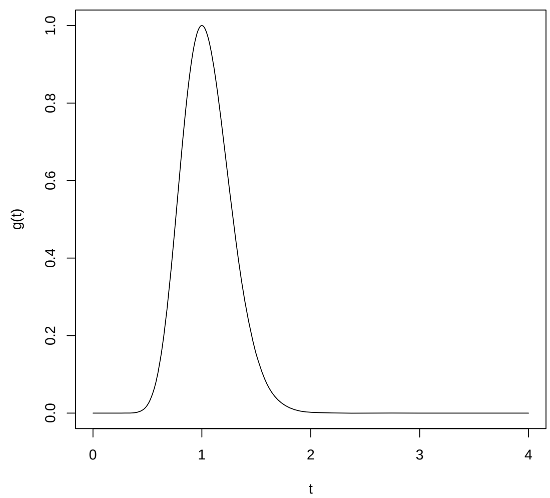

Figure 2: Sketch of graph of $g(t)$ for $0 \leqslant t \leqslant 4$

there exist positive numbers t_1 and t_2 such that that $t_1 < 1 < t_2$ and

$$\Lambda \leqslant c \quad \text{if and only if} \quad t \leqslant t_1 \quad \text{or} \quad t \geqslant t_2.$$

Thus, the LRT rejects H_o iff

$$\frac{\widehat{\theta}}{\theta_o} \leqslant t_1 \text{ or } \frac{\widehat{\theta}}{\theta_o} \geqslant t_2 \quad \text{for some } 0 < t_1 < 1 < t_2, \tag{1}$$

where $\widehat{\theta} = \dfrac{1}{4n} \displaystyle\sum_{i=1}^{n} X_i$. Thus, the LRT for

$$\mathrm{H}_o\colon \quad \theta = \theta_o \quad \text{versus} \quad \mathrm{H}_1\colon \quad \theta \neq \theta_o$$

may be based upon the statistic

$$W = \sum_{i=1}^{n} X_i.$$

In fact, we see from (1) that the rejection region for the LRT is

$$R\colon \quad \frac{2W}{\theta_o} \leqslant 8nt_1 \quad \text{or} \quad \frac{2W}{\theta_o} \geqslant 8nt_2,$$

or

$$R: \quad W \leqslant c_1 \quad \text{or} \quad W \geqslant c_2,$$

for some positive constants, c_1 and c_2, with $c_1 < c_2$.
It is possible to show that, if the null hypothesis is true, then the moment generating function of $\dfrac{2W}{\theta_o}$ is

$$\left(\frac{1}{1-2t}\right)^{4n},$$

which is the mgf of a χ^2 random variable with $8n$ degrees of freedom. It then follows that

$$\frac{2W}{\theta_o} \sim \chi^2(8n),$$

if H_o is true. □

(b) For $\theta_o = 4$ and $n = 5$, find c_1 and c_2 so that the test rejects H_o when $W \leqslant c_1$ or $W \geqslant c_2$ has a significance level $\alpha = 0.05$.

Solution: If H_o is true, then $\dfrac{2W}{\theta_o} \sim \chi^2(40)$. We first find positive numbers, a_1 and a_2, with $a_1 < a_2$, and

$$F_{_{2W/\theta_o}}(a_1) = 0.025 \quad \text{and} \quad F_{_{2W/\theta_o}}(a_2) = 0.975.$$

this yields

$$a_1 \approx 24.43 \quad \text{and } a_2 \approx 59.34.$$

Setting

$$c_1 = \frac{a_1\theta_o}{2} \approx 48.86 \quad \text{and } c_2 = \frac{a_2\theta_o}{2} \approx 118.68,$$

it follows that the LRT which rejects H_o if

$$W \leqslant c_1 \quad \text{or} \quad W \geqslant c_2$$

has significance level $\alpha = 0.05$. □

7. Suppose that $X_1, X_2, \ldots, X_n$ form a random sample from a normal$(0, \sigma^2)$ distribution. We wish to test

$$\mathrm{H}_o\colon \quad \sigma^2 \leqslant 2 \quad \text{versus} \quad \mathrm{H}_1\colon \quad \sigma^2 > 2.$$

(a) Show that there exists a uniformly most powerful (UMP) test at every significance level α.

Solution: We first consider the test of simple hypotheses

$$\mathrm{H}_o\colon \quad \sigma^2 = 2 \quad \text{versus} \quad \mathrm{H}_1\colon \quad \sigma^2 = \sigma_1^2, \tag{2}$$

where $\sigma_1^2 > 2$.
The likelihood function for this situation is

$$L(\sigma^2 \mid x_1, x_2, \ldots, x_2) = \frac{1}{(2\pi)^{n/2}\ \sigma^n}\ e^{-\sum_{i=1}^n x_i^2/2\sigma^2},$$

thus, the likelihood ratio statistic for the test of simple hypotheses in (2) is

$$\Lambda(x_1, x_2, \ldots, x_n) = \left(\frac{\sigma_1^2}{2}\right)^{n/2}\ e^{-\frac{w}{4}\left(1 - \frac{1}{\sigma^2/2}\right)},$$

where we have written

$$w = \sum_{i=1}^n x_i^2.$$

Setting $a = \dfrac{\sigma_1^2}{2}$ and $b = 1 - \dfrac{1}{a}$, we see that the likelihood ratio statistic is a function of w; more precisely,

$$\Lambda(x_1, x_2, \ldots, x_n) = g(w),$$

where

$$g(w) = a^{n/2} e^{-bw/4}.$$

Since $\sigma_1^2 > 2$, $a > 1$ and therefore $b > 0$; so that, $g(w)$ decreases to 0 as w increases to infinity. Thus, for any $c \in (0,1)$ we can find a positive, w_1, large enough so that $g(w_1) = c$. Furthermore,

$$g(w) \leqslant c \quad \text{for all} \quad w \geqslant w_1,$$

since $g(w)$ decreases with increasing w. Thus, the LRT rejection region for the simple hypotheses in (2) is equivalent to the region

$$R\colon \quad W \geqslant w_1.$$

Observe that, if H_o in (2) is true, then the statistic $\dfrac{1}{2}W$ has a $\chi^2(n)$ distribution, since $X_1, X_2, \ldots, X_n$ are iid normal$(0,2)$ random variables when H_o is true. Thus, given $\alpha \in (0,1)$, let

$$w_\alpha = F^{-1}_{W/2}(1-\alpha).$$

Then, the test that rejects H_o if

$$\frac{1}{2}W \geqslant w_\alpha$$

has significance level α. Equivalently, the test that rejects H_o if

$$W \geqslant c_\alpha,$$

where

$$c_\alpha = 2w_\alpha,$$

has significance level α. By the Neyman–Pearson Lemma, this is the most powerful test at level α.
Next, consider the test of simple hypotheses

$$\mathrm{H}_o\colon \quad \sigma^2 = \sigma_o^2 \quad \text{versus} \quad \mathrm{H}_1\colon \quad \sigma^2 = \sigma_1^2, \tag{3}$$

where $\sigma_o^2 \leqslant 2$ and $\sigma_1 2 > 2$. Then, the likelihood ratio statistic for this test is

$$\Lambda(x_1, x_2, \ldots, x_n) = \left(\frac{\sigma_1^2}{\sigma_o^2}\right)^{n/2} e^{-\frac{w}{2\sigma_o^2}\left(1-\frac{1}{\sigma^2/\sigma_o^2}\right)},$$

Thus, since $\sigma_1^2 > \sigma_o^2$, we see that Λ is a decreasing function of w, Consequently, the test that rejects H_o if

$$W \geqslant c_\alpha,$$

is the most powerful at level α. Since this is the case for all $\sigma_o^2 \leqslant 2$ and $\sigma_1^2 > 2$, this test is the uniformly most powerful test at significance level α. □

(b) Show that the UMP test found in part (a) rejects H_o when

$$\sum_{i=1}^{n} X_i^2 \geqslant c,$$

for some $c > 0$, and determine the value of c so that the significance level of the test is $\alpha = 0.05$.

Solution: We solved the first part of this problem in part (a) for $c = c_\alpha$.

To determine c_α for $\alpha = 0.05$ solve the equation

$$F_{W/2}(w_\alpha) = 1 - \alpha,$$

where $\frac{1}{2}W \sim \chi^2(n)$. Then, $c_\alpha = 2w_\alpha$. □

8. In a given city, it is assumed that the number of automobile accidents in a given year follows a Poisson distribution. Suppose that it is known that, in past years, the average number of accidents per years was 15. Suppose that this year the number of accidents has been 10. Is it justified to claim that the rate of accidents has dropped?

   To answer this question, set up an appropriate hypothesis test. State your assumptions clearly and justify your conclusions.

Solution: To answer this question, we will make an inference based on a single observation, Y, where $Y \sim \text{Poisson}(\lambda)$.

We will test the hypothesis

$$\text{H}_o: \quad \lambda = \lambda_o,$$

where $\lambda_o = 15$, versus the alternative

$$\text{H}_1: \quad \lambda < \lambda_o.$$

We use a likelihood ratio test. In order to do this, we firs note that the likelihood function in this case is

$$L(\lambda \mid y) = \frac{\lambda^y}{y!}\ e^{-\lambda}$$

The maximum likelihood occurs when $\lambda = y$; thus, we set $\widehat{\lambda} = y$.

To compute the likelihood statistic,

$$\Lambda(y) = \frac{L(\lambda_o \mid y)}{\sup\limits_{\lambda<\lambda_o} L(\lambda \mid y)},$$

We consider two cases:

(i) $\widehat{\lambda} < \lambda_o$;

(ii) $\widehat{\lambda} \geqslant \lambda_o$.

In case (i), $\sup\limits_{\lambda<\lambda_o} L(\lambda \mid y) = L(\widehat{\lambda} \mid y)$, so that $\Lambda(y) = \dfrac{L(\lambda_o \mid y)}{L(\widehat{\lambda} \mid y)}$.

In case (ii), $\sup\limits_{\lambda<\lambda_o} L(\lambda \mid y) = L(\lambda_o \mid y)$, so that $\Lambda(y) = 1$.

We then have that

$$\Lambda(y) = \begin{cases} \dfrac{\lambda_o^y\, e^{-\lambda_o}}{\widehat{\lambda}^y\, e^{-\widehat{\lambda}}} & \text{if } \widehat{\lambda} < \lambda_o; \\ \\ 1 & \text{if } \widehat{\lambda} \geqslant \lambda_o, \end{cases}$$

which we can write in terms of $t = \dfrac{\widehat{\lambda}}{\lambda_o}$ as $\Lambda(y) = g(t)$, where

$$g(t) = \begin{cases} \dfrac{1}{t^{\lambda_o t}\, e^{-\lambda_o(t-1)}} & \text{if } t < 1; \\ \\ 1 & \text{if } t \geqslant 1. \end{cases}$$

Observe that the function $g(t)$ increases strictly from $e^{-\lambda_o}$ to 1 for $0 < t < 1$. Consequently, for any $c < 1$, with $c > e^{-\lambda_o}$, there exists $t_1 < 1$ such that $g(t_1) = c$, and

$$g(t) \leqslant c \quad \text{for all } t \leqslant t_1.$$

Consequently, the LRT rejection region

$$R\colon \quad \Lambda \leqslant c,$$

is equivalent to

$$R\colon \quad \frac{\widehat{\lambda}}{\lambda_o} \leqslant t_1,$$

or

$$R\colon \quad Y \leqslant t_1\lambda_o \equiv a,$$

where $a < \lambda_o$, and $Y \sim \text{Poisson}(\lambda)$.

Using the test statistic Y, and the observed value $y = 10$, we may compute the p–value

$$p\text{–value} = \mathrm{P}(Y \leqslant 10), \quad \text{where } Y \sim \text{Poisson}(\lambda_o).$$

in the case $\lambda_o = 15$, we obtain that

$$p\text{–value} \approx 0.1185,$$

thus, we cannot reject H_o at any significant level $\alpha \leqslant 10\%$. Hence, we cannot conclude that the rate of accidents has dropped at a significance level of less than 10%. $\square$